Exhibit 4

Journal of Economic Dynamics & Control 36 (2012) 501–522

Contents lists available at SciVerse ScienceDirect

Journal of Economic Dynamics & Control

journal homepage: www.elsevier.com/locate/jedc

The market impact of a limit order

Nikolaus Hautsch [a,b,c,*], Ruihong Huang [a]

[a] Institute for Statistics and Econometrics, Humboldt-Universität zu Berlin, Spandauer Str. 1, 10178 Berlin, Germany
[b] Center for Applied Statistics and Economics (CASE), Humboldt-Universität zu Berlin, Germany
[c] Center for Financial Studies, Frankfurt

ARTICLE INFO

Article history:
Received 23 July 2010
Accepted 19 September 2011
Available online 17 December 2011

JEL classification:
G14
G32
G17

Keywords:
Price impact
Limit order
Impulse response function
Cointegration

ABSTRACT

We quantify the short-run and long-run price effect of posting a limit order in an order book market by proposing a high-frequency cointegrated VAR model for quotes and order book depth. Estimating impulse response functions based on data from 30 stocks traded at Euronext Amsterdam we show that limit orders have significant market impacts. The strength and direction of quote responses depend on the incoming orders' aggressiveness, their size and the state of the book. The effects are qualitatively stable across the market. Cross-sectional variations in the magnitudes of price impacts are well explained by the underlying trading frequency and relative tick size.

1. Introduction

It is well known that the revelation of trading intention adversely affects asset prices. As also confirmed by theoretical studies,[1] passive order placement through limit orders incurs significant market impact even if the order is not been executed. In financial practice, the risk to 'scare' and to ultimately shift the market by limit order placements is well-known and is taken into account in trading strategies. As a consequence, liquidity provision through hidden orders, which allow traders to partly (or entirely) conceal order volume, has gained popularity. However, despite of the importance of limit order strategies in modern trading, the actual impact of an incoming (visible) limit order on the subsequent price process is still hardly explored and quantified. In fact, while the analysis of the price impact resulting from a trade is a classical topic in traditional market microstructure research (see, e.g., Hasbrouck, 1991; Dufour and Engle, 2000; Engle and Patton, 2004), empirical evidence on the market impact of limit order placements is addressed by only few recent studies as Eisler et al. (in press) and Cont et al. (2011).

This paper aims at filling this gap in the literature and addresses the following empirical research questions: (i) How strong is the short-run and long-run impact of an incoming limit order in dependence of its position in the book, its size and the state of the book? (ii) Are ask and bid quote responses to incoming limit orders widely symmetric or is there evidence for an asymmetric re-balancing of the book? (iii) How different is the market impact of a limit order compared to that caused by a trade of similar size? (iv) How stable are these effects across the market and do they depend on stock-specific characteristics, such as the underlying trading intensity, minimum tick size and average trade size?

* Corresponding author at: Institute for Statistics and Econometrics, Humboldt-Universität zu Berlin, Spandauer Str. 1, 10178 Berlin, Germany.
E-mail addresses: nikolaus.hautsch@wiwi.hu-berlin.de (N. Hautsch), ruihong.huang@wiwi.hu-berlin.de (R. Huang).

[1] See, e.g., Parlour and Seppi (2008), Boulatov and George (2008) and Rosu (2010).

0165-1889/$ - see front matter © 2012 Published by Elsevier B.V.
doi:10.1016/j.jedc.2011.09.012

We propose modelling the processes of ask and bid quotes as well as several levels of depth volume on both sides of the market in terms of a cointegrated vector autoregressive (VAR) model. This framework allows studying the price impact of limit orders by means of impulse response functions. Each limit order is represented by a shock disturbing the multivariate system of quotes and depths and influencing it dynamically over time. Designing the shock vectors in a specific way allows us to characterize the type of the limit order represented by its size and its position in the order queue as well as the current state of the book.

The motivation for using a cointegrating system stems from the fact that ask and bid quotes are naturally integrated and tend to move in locksteps. Cointegration analysis reveals a stationary linear combination of bid and ask quotes which closely resembles the bid–ask spread. The idea of jointly modelling ask and bid quote dynamics in terms of a cointegrated system originates from Engle and Patton (2004) based on the work of Hasbrouck (1991) and has been used in other approaches, such as Hansen and Lunde (2006) and Escribano and Pascual (2006). Our setting extends and modifies this approach in two major directions: Firstly, we model quotes and depth simultaneously. This yields a novel type of order book model capturing not only quote and depth dynamics but implicitly also dynamics of midquotes, midquote returns, spreads, spread changes as well as order book imbalances. Secondly, we model the system not only on a trade-to-trade basis but exploit the complete order arrival process. Therefore, the model captures all relevant trading characteristics in a limit order book market and thus provides a complete description of the order book in a range close to the best quotes. Hence, the model is particularly useful for liquid assets where most of the market activity is concentrated at the best quote levels. In this sense, the approach complements dynamic models for order book curves such as proposed by Härdle et al. (2009) and Russell and Kim (2010).

The proposed quote and depth model is estimated by Johansen's (1991) full information maximum likelihood estimator using high-frequency order book data for 30 stocks traded on Euronext Amsterdam covering a sample period over two months in 2008. We find strong evidence for the existence of common stochastic components in quotes and corresponding depths resulting in cointegration relations which significantly deviate from the bid–ask spread. In this sense, our results shed some light on the strength of co-movements in ask and bid prices depending on the underlying depth. Indeed, it turns out that order book inventory is highly persistent and reveals high-frequency dynamics resembling (near-)unit-root behavior. We show that incoming limit orders have significant impacts on subsequent ask and bid processes. It turns out that the magnitude and direction of quote adjustments strongly depend on the order's aggressiveness, its (relative) size and the prevailing depth in the book. In particular, we show the following results: (i) Quote adjustments are the stronger and the faster, the closer the incoming order is posted to the market. Most significant effects are reported for orders posted on up to two levels behind the market. For less aggressive orders, virtually no effects can be quantified. (ii) Limit orders temporarily narrow the spread. Converse effects are shown for market orders. In the long-run, these effects are reverted back in an asymmetric way. (iii) Large limit orders posted inside of the spread induce severe long-run effects pushing the market in the intended trading direction. In contrast, small limit orders posted inside of the spread tend to be picked up quickly inducing adverse price reactions. (iv) The long run market impact of aggressive market orders walking through the book is the higher the smaller the prevailing depth behind the market. (v) The effects are qualitatively stable across the market, where the absolute magnitudes of price impacts differ in dependence of underlying stock-specific characteristics. It turns out that approximately 60–80% of the cross-sectional variation in market impacts can be explained by the trading frequency and the minimum tick size.

The remainder of this paper is structured as follows. In Section 2, we describe the trading structure of Euronext Amsterdam and provide descriptive statistics. The econometric approach is explained in Section 3. Section 4 gives the estimation results and Section 5 provides the quantified price impacts of different types of limit orders. Finally, Section 6 concludes.

2. Data and market environment

Euronext is a purely electric limit order book market with price and time order precedence. During the continuous trading period between 9:00 and 17:30 CET, limit orders are submitted to a centralized computer system where they are matched to prevailing standing limit orders on the opposite side. If there is no match or the matched volume in the system is insufficient to exhaust the incoming order, the remaining order volume is placed in the order book. Euronext supports various order types like pure market orders (immediate order execution without a price limit), stop orders (automatic issuing of limit orders or market orders when a given price is reached), fill-or-kill (FOK) orders or iceberg orders.

Our dataset comprises limit order book (LOB) data of the 30 most frequently traded stocks at Euronext Amsterdam between August 1 and September 30, 2008. Since on September 1st, Euronext changed the minimum tick size for some stocks, we analyze the two months August and September separately. This allows us to study the robustness of our findings under changing market conditions. Since these two months represent a generally turbulent market period, we further robustify our findings by replicating our analysis for a period which is less volatile. As we obtain quantitatively similar results, our findings can be seen as representative for different market conditions.[2]

[2] These results are not shown in the paper but are available on our web appendix or upon request.

The data contains information on the prevailing market depth (in terms of the number of shares) for the five best quotes on both sides of the market. Every trade and change of the order book is recorded in milliseconds. Preliminary analyses (which are also supported by the findings given in Section 5) show that aggressive limit orders placed close to the best ask and bid have the highest market impact while induced price effects significantly decline with the distance to the spread. Accordingly, we focus only on the best three price levels in the book. Unlike the trade data which is well filtered by built-in filters in the database,[3] the order book data is completely raw. We remove observations where (i) the spread is zero or negative, and (ii) ask or bid quotes change by more than 2%.[4] Moreover, to remove effects due to the opening and closing of the market, we discard data of the first five and last five minutes of the continuous trading period.

Matching of trade and LOB data is achieved by a matching algorithm which is described in detail in Appendix A. This algorithm matches a trade with the corresponding LOB observation by comparing its price and volume with the resulting changes of quotes and depths in the book within an adaptively chosen time window. It minimizes the probability of misclassifications and, as a by-product, provides an estimate of the time asynchronicity between trade and LOB records.[5] To classify the initiation type of trades, we use a hybrid procedure according to Lee and Ready (1991). Firstly, we determine the type of trades which are located in more than 1 s time distance to previous trades using the mid-quote method. I.e., if a trade occurs with a price greater (less) than the most current mid-quote, it is classified as a buy (sell). If the trade price equals the mid-quote, it is marked as 'undetermined'. Secondly, 'undetermined' trades and trades which follow previous trades in less than 1 s time distance are classified by the tick-test method. Accordingly, if the trade price is higher (lower) than the previous one, it is identified as a buy (sell). If it does not change the price, it is categorized as the same type as the previous one. Finally, we identify sub-trades arising from the execution of a big market order against several (smaller) limit orders if they occur in less than one second after the previous trade and have the same initiation types. All corresponding sub-trades are consolidated to a single trade.

Table 1 gives descriptive statistics of the resulting August data used in the paper.[6] We observe significantly more limit order activities than market orders. The average bid–ask spread is decreasing with the liquidity of the underlying stock. On average, second level market depth is higher than first level depth while it is approximately equal to the depth on the third level.

3. Econometric modelling

3.1. A cointegrated VAR model for quotes and depths

Denote t as a (business) time index, indicating all order book activities, i.e., incoming limit or market orders as well as limit order cancellations. Then, p_t^a and p_t^b denote the best log ask and bid quotes instantaneously after the t-th order activity and $v_t^{a,j}$ and $v_t^{b,j}$ for $j=1,\ldots,k$, define the log depth on the j-th best observed quote level on the ask and bid side, respectively. Furthermore, we introduce two dummy variables, BUY_t and $SELL_t$ indicating the occurrence of buy and sell trades, respectively. The inclusion of these two variables is necessary to distinguish between the effects caused by a market order and that induced by a cancellation. Both events remove volume from the book, however, presumably have quite different long run market impacts. Table 2 gives a detailed description of the variables.

To capture the high-frequency dynamics in quotes and depths, we define a $K=(4+2\times k)$-dimensional vector of endogenous variables

$$y_t \coloneqq [p_t^a, p_t^b, v_t^{a,1}, \ldots, v_t^{a,k}, v^{b,1}, \ldots, v_t^{b,k}, BUY_t, SELL_t]'.$$

Note that the quote levels associated with $v_t^{a,j}$ and $v_t^{b,j}$ are not observed on a *fixed* grid at and behind the best quotes. Hence, their price distance to p_t^a and p_t^b is not necessarily exactly $j-1$ ticks but might be higher if there are no limit orders on all possible intermediate price levels behind the market. To capture such 'gaps' in the order book, we could also include the limit prices associated with each order level posted behind the market and thus correspondingly extend the vector y_t. However, we decided to disregard this information because of two reasons. Firstly, Hautsch and Huang (2011) show that trades "walking through the book", i.e., trades absorbing more than one price level in the limit order book occur extremely rarely for liquid stocks. Secondly, in liquid markets, the tick levels close to the best quotes are indeed mostly filled such that limit prices are on a fixed grid with constant distance to the corresponding best quotes. Consequently, we expect that the inclusion of all individual limit prices does not provide any additional information but just increases the dimension of the system. Finally, modelling log volumes instead of plain volumes is a common practice in many empirical studies to reduce the impact of extraordinarily large volumes. This is also suggested by Potters and Bouchaud (2003) studying the

[3] Besides recording errors, block trades and trades in auction periods are excluded.

[4] In order to limit the volatility, Euronext NSC suspends continuous trading if prices change by more than 2%. This is not exactly the same rule as that implemented here, but it is reasonably mimicked.

[5] Due to technological progress in the last decades, the time delay between trade and quote records is nowadays hardly greater than 1 s. Consequently, the '5-s' rule according to Lee and Ready (1991), which has been commonly used in empirical market microstructure literature, is not appropriate anymore for more recent datasets.

[6] Due to the aforementioned change in the minimum tick size, it is not appropriate to present joint summary statistics for both months. However, as the descriptive statistics for September are very similar to that for August, we do not present them here.

Table 1
Summary of synchronized trade and order book data. The sample consists of the 30 most frequently traded stocks on Euronext Amsterdam. Market depth is measured in thousand shares. L1–L3 denote the order book level one to three. The period is from 1 to 31 August 2008.

Stocks	#trades per day	#LO activ. per day	Ask			Bid			Mean of ask depth			Mean of bid depth		
			Min	Mean	Max	Min	Mean	Max	L1	L2	L3	L1	L2	L3
ING	1606.8	66 569.1	20.255	21.518	23.290	20.250	21.507	23.275	3.64	3.94	4.12	3.45	3.90	4.14
FOR	1304.6	27 574.0	8.770	9.351	10.160	8.760	9.338	10.150	16.78	25.76	25.03	16.35	26.25	24.20
RDSa	1166.2	48 630.6	21.900	22.991	23.935	21.890	22.981	23.930	4.30	5.21	5.80	4.00	5.06	5.59
UNc	1152.1	46 023.7	17.110	18.635	19.670	17.100	18.625	19.660	4.76	5.24	6.44	4.52	5.33	6.49
AHLN	1119.4	18 730.3	7.540	8.510	8.970	7.530	8.502	8.960	7.89	9.80	10.23	8.18	10.64	10.59
PHG	1108.3	34 722.0	20.875	22.381	23.465	20.870	22.368	23.450	2.18	2.36	2.70	1.95	2.19	2.59
AEGN	982.5	43 270.2	7.290	7.909	8.400	7.280	7.902	8.395	5.12	4.99	4.86	4.98	4.98	4.79
AKZO	960.0	20 061.2	35.460	39.571	41.920	35.400	39.541	41.910	0.89	0.96	1.00	0.78	0.90	0.98
KPN	954.0	20 733.8	10.915	11.274	11.680	10.905	11.266	11.670	9.61	12.10	12.77	8.79	10.57	11.57
TNT	949.7	20 412.7	22.040	24.598	27.000	22.030	24.566	26.970	1.57	1.91	2.15	1.51	1.96	2.24
HEIN	927.2	19 782.1	29.540	31.796	33.660	29.520	31.767	33.600	0.98	1.10	1.13	0.92	1.00	1.04
ISPA	903.1	35 708.2	49.990	52.694	56.440	49.910	52.661	56.420	1.85	2.76	3.66	1.97	3.08	3.84
ASML	853.8	26 249.5	14.290	15.964	17.400	14.280	15.949	17.390	3.80	5.86	6.50	3.48	5.21	6.01
DSMN	826.7	21 574.5	36.050	37.919	40.000	36.020	37.886	39.990	0.77	0.87	0.99	0.77	0.88	0.99
SBMO	603.7	18 676.3	13.530	14.934	16.700	13.520	14.911	16.680	1.84	2.63	2.99	1.76	2.51	2.79
TOM2	505.3	16 822.0	14.340	16.017	17.550	14.300	15.987	17.540	1.31	1.71	2.06	1.25	1.69	1.75
FUGRc	505.0	8846.5	43.620	47.701	53.200	43.610	47.631	53.180	0.56	0.54	0.52	0.49	0.49	0.47
WLSNc	548.8	16 003.6	14.610	15.973	17.020	14.550	15.950	17.000	1.92	1.88	1.96	1.94	1.83	1.89
RAND	543.4	17 265.2	17.710	19.432	21.430	17.690	19.397	21.400	1.09	1.56	1.75	1.07	1.47	1.47
ELSN	488.5	29 702.2	10.390	11.049	11.510	10.350	11.035	11.500	7.27	11.57	11.96	6.81	11.34	12.44
BOSN	419.6	8013.0	32.320	36.323	41.900	32.250	36.247	41.890	0.52	0.52	0.49	0.53	0.51	0.47
BAMN	416.8	6334.1	9.900	10.736	12.220	9.860	10.714	12.200	2.06	2.35	2.38	1.99	2.25	2.19
SR	347.5	6396.6	10.370	11.588	13.200	10.360	11.563	13.180	1.70	1.80	1.76	1.72	1.71	1.48
CSMNc	340.2	7478.4	17.910	20.395	24.260	17.890	20.361	24.240	0.81	0.88	0.92	0.84	0.90	0.91
COR	327.1	12 103.2	47.090	49.273	51.210	47.010	49.175	51.140	0.43	0.41	0.37	0.39	0.38	0.34
IMUN	292.7	5735.9	14.300	16.178	17.710	14.280	16.148	17.700	0.92	1.17	1.24	0.85	0.91	0.88
SMTNc	272.4	7648.8	43.920	52.282	60.440	43.840	52.112	60.300	0.22	0.25	0.22	0.26	0.27	0.26
NUTR	256.6	8043.2	41.160	43.275	44.900	41.120	43.192	44.890	0.40	0.36	0.33	0.37	0.38	0.38
USGP	248.5	6342.3	9.670	11.198	12.630	9.650	11.168	12.600	1.47	1.51	1.41	1.59	1.39	1.19
HEIO	181.0	14 011.0	27.120	29.854	31.300	27.080	29.809	31.290	0.44	0.53	0.61	0.50	0.64	0.70

Table 2
Variable definitions. Events include limit order submissions, executions and cancellations. The market depth refers to the pending volume at the ordered available price levels in the LOB.

Variable	Description
p_t^a	Logarithm of the best ask after the t-th event.
p_t^b	Logarithm of the best bid after the t-th event.
$v_t^{a,l}$	Logarithm of market depth at the l-th best ask after the t-th event.
$v_t^{b,l}$	Logarithm of market depth at the l-th best bid after the t-th event.
BUY_t	Dummy equal to one if the t-th event is a buyer-initiated trade.
$SELL_t$	Dummy equal to one if the t-th event is a seller-initiated trade.

statistical properties of market impacts of trades. Moreover, using logs implies that changes in market depth can be interpreted as *relative* changes with respect to the current depth level.

Hence, we model log quotes, log depths and trading indicators as a restricted cointegrated vector autoregressive model of the order p (VAR(p)) with the vector error correction (VEC) form

$$\Delta y_t = \mu + \alpha\beta' y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + u_t, \tag{1}$$

where u_t is white noise with covariance matrix Σ_u, μ is a constant, Γ_i with $i = 1, \ldots, p-1$ is a $K \times K$ parameter matrix, α and β denote the $K \times r$ loading and cointegrating matrices with $r < K$. As we can safely assume that the trading indicators BUY_t and $SELL_t$ are stationary, we restrict the two first columns of β to $\beta_1 = [0, \ldots, 0, 1, 0]'$ and $\beta_2 = [0, \ldots, 0, 0, 1]'$.

For the impulse-response analysis below, it turns out to be more convenient to work with the reduced VAR representation in terms of the level of y_t,

$$y_t = \mu + \sum_{i=1}^{p} A_i y_{t-i} + u_t, \tag{2}$$

where $A_1 := I_K + \alpha\beta' + \Gamma_1$ with I_K denoting a $K \times K$ identity matrix, $A_i := \Gamma_i - \Gamma_{i-1}$ with $1 < i < p$ and $A_p := -\Gamma_{p-1}$.

We estimate the model (1) employing the Full Information Maximum Likelihood (FIML) estimator proposed by Johansen (1991) and Johansen and Juselius (1990). Then, following Lütkepohl and Reimers (1992), we transform these estimates to representation (2). The corresponding procedure is shown in Appendixes B and C. By imposing the stationarity restrictions β_1 and β_2, all elements in the other cointegrating vectors associated with BUY_t and $SELL_t$ are automatically set to zero. This is guaranteed by the orthogonality among the estimated cointegrating vectors implied by FIML.

Note that market depth enters the vector y_t in levels and thus is treated as a possibly non-stationary variable. Though this is counter-intuitive for the behavior of depth over longer horizons, it is a reasonable assumption if depth is observed on very high frequencies. Moreover, modelling both quotes *and* depth in terms of a cointegration system guarantees consistency of parameter estimates irrespective of the possible (non-)stationarity of order book depth. Even if depth is truly stationary (and thus just corresponds to a (spurious) cointegration relation for itself), FIML estimates are consistent (though obviously not efficient).[7] Since we employ a high number of observations, the possible loss of efficiency due to the neglect of a (stationarity) restriction is not very harmful in our context. If, however, we impose stationarity of depth and correspondingly restrict the cointegration vectors, we run the risk of producing inconsistent estimates if the restriction does not hold. Indeed, unit root tests applied in Section 4.1 indicate that the assumption of a unit root in depth observed on high frequencies cannot be rejected for many stocks. These arguments support the usefulness of a more robust statistical inference in the form of an unrestricted cointegration system.

Model (2) can be further rotated in order to represent dynamics in spreads, relative spread changes, midquotes, midquote returns as well as (ask–bid) depth imbalances. Hence, the model is sufficiently flexible to capture the high-frequency dynamics of all relevant trading variables.[8]

Finally, in models involving only quote dynamics (e.g., Engle and Patton, 2004) or spread dynamics (e.g., Lo and Sapp, 2006), the error correction term $\beta' y_t$ is typically assumed to be equal to the spread implying a linear restriction $R'\beta = 0$ with $R' = [1, 1, 0, \ldots, 0]$. However, given the potential non-stationarity of order book depth, we do *not* impose this assumption here. As depth might contain information on the equilibrium (long run) state of the order book as well, we expect the existence of cointegration relations differing from spreads and involving both quotes *and* depths. As shown in the remainder of the paper, this notion is actually supported by the data.

3.2. Limit orders as shocks to the system

In this section, we illustrate how to represent incoming orders as shocks to the system specified in Eq. (2). Whenever an order enters the order book, it (i) will change the depth in the book, (ii) may change the best quotes depending on which position in the queue it is placed, and (iii) will change the trading indicator dummy in case of a market order. We represent these changes in terms of an impulse vector $\delta := [\delta_v', \delta_p', \delta_d']'$ with δ_v being a $2k \times 1$ vector associated with shocks to depth, δ_p denoting a 2×1 vector consisting of shocks to the quotes and δ_d being a 2×1 vector representing shocks to the trading indicator dummy.

We design impulse response vectors associated with five scenarios commonly faced by market participants. As graphically illustrated by Figs. 1–4, a three-level order book is initialized by the best ask $p_t^a = 1002$, best bid $p_t^b = 1000$, second best ask 1003, second best bid 999, and levels of depths on the bid side $V_t^{b,1} = 1$, $V_t^{b,2} = 1.5$, $V_t^{b,3} = V_t^{b,4} = 1.4$. The following scenarios are considered[9]:

Scenario 1a (*Normal limit order*). Arrival of a buy limit order with price 1000 and size 0.5 to be placed *at the market*. As shown in Fig. 1, this order will be consolidated at the best bid without changing the prevailing quotes. Because the initial depth on the first level is 1.0, the change of the log depth is $\ln(1.5) \approx 0.4$. Correspondingly, the shock vectors are given by $\delta_v = [0, 0, 0, 0.4, 0, 0]'$, $\delta_p = \delta_d = [0, 0]'$.

Scenario 1b (*Passive limit order*). Arrival of a buy limit order with price 999 and size 0.5 to be posted *behind the market*. As in the scenario above, it does not change the prevailing quotes and only affects the depth at the second best bid. We have $\delta_v = [0, 0, 0, 0, \ln(2) - \ln(1.5) \approx 0.29, 0]'$, $\delta_p = \delta_d = [0, 0]'$.

[7] See, for instance, Example 3.1 in Johansen (1995) for an illustration of this argument.

[8] Note that we do not impose an explicit constraint ensuring the positiveness of bid–ask spreads. As shown on the companion website, this restriction is implicitly satisfied by our estimates in virtually all cases.

[9] For the sake of brevity, the scenarios are only characterized for buy orders. For sell orders, the setting is correspondingly adapted to the other side of the market.

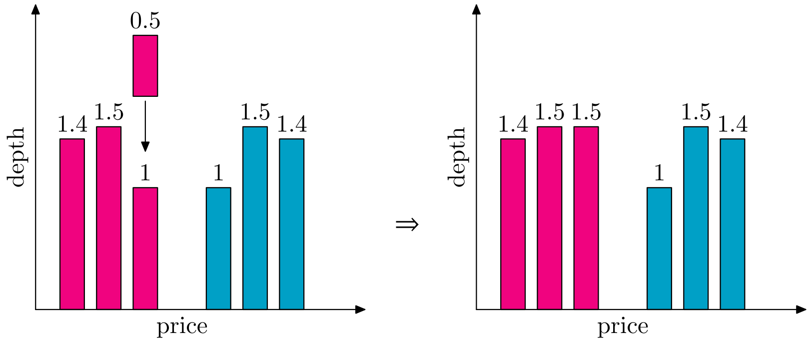

Fig. 1. (*Scenario* 1a (*normal limit order*)): an incoming buy limit order with price 1000 and size 0.5. It affects only the depth at the best bid without changing the prevailing quotes or resulting in a trade.

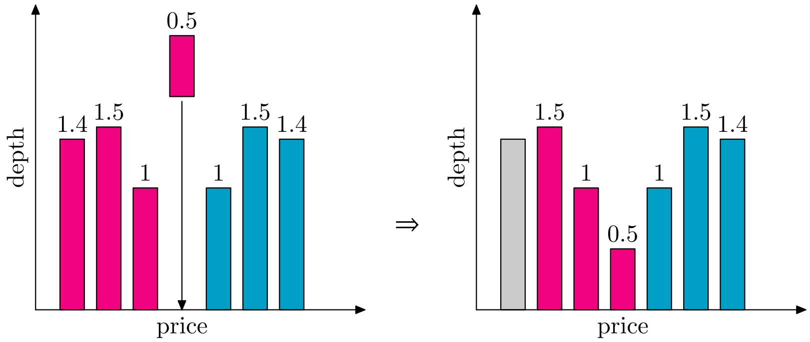

Fig. 2. (*Scenario* 2 (*aggressive limit order*)): an incoming buy limit order with price 1001 and size 0.5 improving the best bid and changing all depth levels on the bid side of the order book.

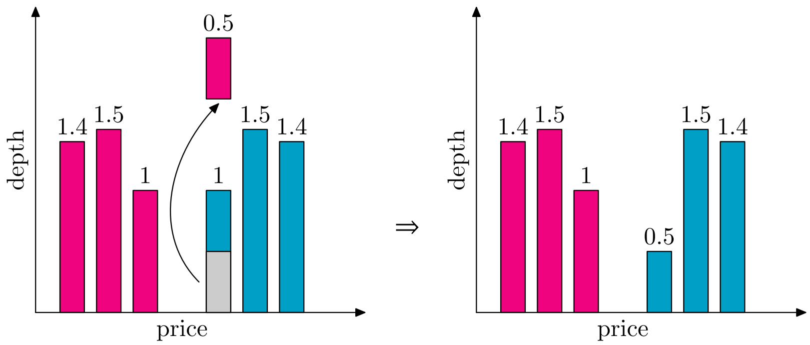

Fig. 3. (*Scenario* 3 (*normal market order*)): an incoming buy market order with price 1002 and size 0.5 which results in a buyer-initiated (BUY) trade.

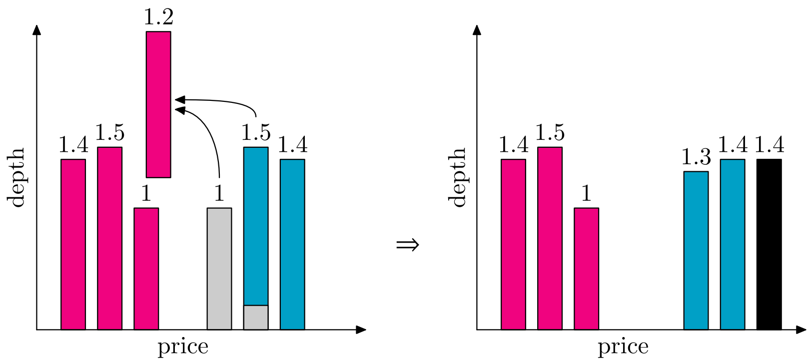

Fig. 4. (*Scenario* 4 (*aggressive market order*)): an incoming buy market order with price 1003 and size 1.2 'walking through' the order book and simultaneously changing all depth levels on the ask side.

Table 3
Shock vectors implied by the underlying five scenarios. Initial order book: best ask $p_t^a = 1002$, best bid $p_t^b = 1000$, second best ask $= 1003$, second best bid $= 999$. Volumes on the ask/bid side: $V_t^{a/b,1} = 1$ at the best bid, $V_t^{a/b,2} = 1.5$ at the second best bid, and $V_t^{a/b,3} = V_t^{a/b,4} = 1.4$ at the third and fourth best bids, respectively. Notation: δ_v denotes shocks on market depths; δ_p denotes shocks on the best ask and bid; δ_d denotes shocks on trading indicator variables.

Scenario	Limit order (dir,price,size)	Shock vectors		
		$\delta_v{}'$	$\delta_p{}'$	$\delta_d{}'$
'Normal limit order'	(Bid,1000,0.5)	[0,0,0,0.4,0,0]	[0,0]	[0,0]
'Passive limit order'	(Bid,999,0.5)	[0,0,0,0,0.29,0]	[0,0]	[0,0]
'Aggressive limit order'	(Bid,1001,0.5)	[0,0,0,−0.69,−0.4,0.07]	[0,0.001]	[0,0]
'Normal market order'	(Bid,1002,0.5)	[−0.69,0,0,0,0,0]	[0,0]	[1,0]
'Aggressive market order'	(Bid,1003,1.2)	[0.26,−0.07,0,0,0,0]	[0.001,0]	[1,0]

Scenario 2 (*Aggressive limit order*). Arrival of a buy limit order with price 1001 and size 0.5 to be posted inside of the current spread. Fig. 2 shows that it improves the best bid by 0.1% and accordingly shifts all depth levels on the bid side. The resulting shock vector is given by $\delta_v = [0,0,0,\ln(0.5) \approx -0.69, \ln(1/1.5) \approx -0.4, \ln(1.5/1.4) \approx 0.07]'$, $\delta_p = [0,0.001]'$ and $\delta_d = [0,0]'$.

Scenario 3 (*Normal market order*). Arrival of a buy order with price 1002 and size 0.5. This order will be executed immediately against standing limit orders at the best ask. Because it absorbs liquidity from the book, it shocks the corresponding depth levels negatively. Fig. 3 depicts the corresponding changes of the order book as represented by $\delta_v = [\ln(0.5) \approx -0.69,0,0,0,0,0]'$, $\delta_p = [0,0]'$ and $\delta_d = [1,0]'$.

Scenario 4 (*Aggressive market order*). Arrival of a buy order with price 1003 and size 1.2. It 'walks up' the order book. As shown in Fig. 4, the best ask quote and all depth levels are simultaneously shifted resulting in the shock vector $\delta_v = [\ln(1.3) \approx 0.26, \ln(1.4/1.5) \approx -0.07,0,0,0,0]'$, $\delta_p = [1/1002 \approx 0.001,0]'$ and $\delta_d = [1,0]'$.

Table 3 summarizes the shock vectors implied by the illustrated scenarios.

3.3. *Measuring the market impact*

We quantify the market impact of limit orders as the implied expected short-run and long-run shifts of the ask and bid after their submissions. This reaction is captured by the impulse response function,

$$f(h;\delta_y) = E[y_{t+h}|y_t+\delta_y, y_{t-1},\ldots] - E[y_{t+h}|y_t, y_{t-1},\ldots], \tag{3}$$

where the shock on quotes, depths and trading indicators is denoted by $\delta_y := [\delta_p', \delta_v', \delta_d']'$ and h is the number of periods (measured in 'order event time').

Note that we do not have to orthogonalize the impulse since contemporaneous relationships between quotes and depths are captured by construction of the shock vector. Moreover, our data is based on the arrival time of orders avoiding time aggregation as another source of mutual dependence in high-frequency order book data.

Using impulse-response analysis to retrieve the market impact has two major advantages. First, in contrast to an analysis of estimated VEC coefficients which only reveals the immediate impact, it enables us to examine both long-run and short-run effects. Second, it allows us to straightforwardly quantify the joint effect induced by simultaneous changes of several variables given a certain state of other variables.

We consider two moving average (MA) representations of the cointegrated VAR model. The first one is based on the reduced form given by Eq. (2). This representation allows us to compute the path of the response function over time. The second one is the Granger representation based on the VECM form in Eq. (1) which enables us to explicitly compute the permanent (long-run) response.

We start our discussion with the first MA representation. The companion VAR(1) form of the VAR(p) model in Eq. (2) is given by

$$Y_t = \boldsymbol{\mu} + \mathbf{A} Y_{t-1} + U_t, \tag{4}$$

where

$$\boldsymbol{\mu} := \underbrace{\begin{bmatrix} \mu \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{Kp\times 1}, \quad Y_t := \underbrace{\begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{bmatrix}}_{Kp\times 1}, \quad U_t := \underbrace{\begin{bmatrix} u_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{Kp\times 1}$$

and

$$\mathbf{A} := \underbrace{\begin{bmatrix} A_1 & \cdots & A_{p-1} & A_p \\ I_K & & 0 & 0 \\ & \ddots & \vdots & \vdots \\ 0 & \cdots & I_K & 0 \end{bmatrix}}_{Kp\times Kp}.$$

Successively substituting Y yields

$$Y_t = M_t + \sum_{i=0}^{t-1} \mathbf{A}^i U_{t-i}, \tag{5}$$

where $M_t = \mathbf{A}^t Y_0 + \sum_{i=0}^{t} \mathbf{A}^i \boldsymbol{\mu}$ consists of terms of an initial value and a deterministic trend, which are irrelevant for the impulse-response analysis. Let $J := [I_K : 0 : \cdots : 0]$ be a $K \times Kp$ selection matrix with $JY_t = y_t$. Pre-multiplying J on both sides of Eq. (5) and using $U_t = J'u_t$ gives

$$y_t = JM_t + \sum_{i=0}^{t-1} J\mathbf{A}^i J' u_{t-i}. \tag{6}$$

Then, the impulse-response function according to Eq. (3) can be written as

$$f(h;\delta_y) = J\mathbf{A}^h J'\delta_y. \tag{7}$$

Given the consistent estimator $\hat{a}$ for $a := \text{vec}(A_1,\ldots,A_p)$ in Eq. (2),

$$\sqrt{T}(\hat{a}-a) \xrightarrow{d} \mathcal{N}(0,\Sigma_{\hat{a}}),$$

Lütkepohl (1990) shows that the asymptotic distribution of the impulse-response function is given by

$$\sqrt{T}(\hat{f}-f) \xrightarrow{d} \mathcal{N}(0, G_h \Sigma_{\hat{a}} G_h'), \tag{8}$$

where $G_h := \partial \text{vec}(f)/\partial \text{vec}(A_1,\ldots,A_p)'$. This expression can be explicitly written as

$$G_h = \sum_{i=0}^{h-1} (\delta_y' J(\mathbf{A}')^{h-1-i} \otimes J\mathbf{A}^i J'). \tag{9}$$

In order to compute the long-run effect, we apply Granger's Representation Theorem to model (1) yielding

$$y_t = C\sum_{i=1}^{t}(u_i+\mu) + C_1(L)(u_t+\mu) + V, \tag{10}$$

where

$$C = \beta_\perp \left(\alpha_\perp' \left(I_K - \sum_{i=1}^{p-1} \Gamma_i \right) \beta_\perp \right)^{-1} \alpha_\perp'. \tag{11}$$

Here, L is the lag operator and the power series $C_1(z)$ is convergent for $|z| < 1+\xi$ for some $\xi > 0$. V depends on initial values, such that $\beta' V = 0$. The Granger representation decomposes the cointegrated process into a random walk term (C term), a stationary process (C_1 term) and a deterministic term (V). Because of the convergence of the series $C_1(z)$, the response implied by this sub-process will be zero in the long run. Moreover, the deterministic term V is irrelevant for the impulse response. Therefore, the permanent response of the system is completely determined by the first term. Note that the shock δ_y causes this term changing by $C\delta_y$. Thus, we can express the permanent response as

$$\overline{f}(\delta_y) := \lim_{h\to\infty} f(h;\delta_y) = C\delta_y. \tag{12}$$

Note that given α and β, $\alpha_\perp$ and $\beta_\perp$ are not uniquely identified. However, the right hand side of Eq. (11) is invariant with respect to the choice of these bases. Therefore, $\overline{f}(\delta_y)$ is unique given the parameters and the shock vector in model (1). In practice, estimated responses and their covariances are obtained by replacing the unknown parameters in Eqs. (7), (8) and (12) by their estimates.

4. Estimation results

The underlying order book data contains bid and ask quotes as well as five levels of depth. Preliminary analyses show that the depths on the fourth and fifth levels do not have significant effects on bid and ask quotes. Therefore, in our empirical study, we only use market depths up to the third level. In order to keep the analysis tractable, we reduce the computational burden induced by the high number of observations by separately estimating the model for each of the 43

trading days. This strategy allows us also to address possible structural changes, e.g., due to stock specific news announcements or overnight effects. The market impact is then computed as the monthly average of individual (daily) impulse responses. Likewise, confidence intervals are computed based on daily averages. To account for a structural break due to the change of the tick size for some stocks on September 1, 2008, we treat the two months August and September separately.

For the sake of brevity we refrain from presenting all individual results for the 30 analyzed stocks in this paper. We rather illustrate the analyzed effects for the stock Fortis (FOR in Table 1) in August 2008. Fortis is one of the most actively traded stocks and is representative for a major part of the market. The results for the remaining stocks and the remaining periods are provided in a web appendix on http://amor.cms.hu-berlin.de/~huangrui/project/impact_of_orders. As shown in the web appendix and discussed in more detail in Section 5.5, the effects are qualitatively remarkably similar across the market though the magnitudes of market impacts differ in dependence of underlying stock-specific characteristics.

The empirical analysis employs a VAR(15) specification which is selected based on residual diagnostics and information criteria. Testing for serial correlation using the Ljung–Box test according to Ljung and Box (1978) reveals almost no remaining serial correlation in the residuals for all regressions based on a 1% level using 10 lags. The corresponding statistics are also recorded in the web appendix.

4.1. Statistical properties of market depth

Fig. 5 provides time series plots of depths on the best ask and third best ask level of the order book for a single (though representative) trading day for Fortis. A general finding is that the depth behind the market is typically greater than that at the market. Furthermore, there is evidence for co-movements between the individual depth levels, partially because of the 'shift' effect induced by aggressive orders, e.g., limit orders posted inside of spreads or market orders completely absorbing the best price levels.

Fig. 6 depicts the unconditional distributions and autocorrelation functions of log market depth. We observe that the distributions of depth behind the market are similar, though they are quite different from those at the market. The same pattern is also observed for the autocorrelation functions. These empirical peculiarities are due to the fact that there are obviously more order activities at the market than behind the market. Consequently, market depth is more frequently changed at the best level inducing a lower persistence than at higher levels. This might also explain why the unconditional distribution of depth is more dispersed than that of depth behind the market.

Table 4 shows the results of Said and Dickey's (1984) Augmented Dickey-Fuller (ADF) and Kwiatkowski et al.'s (1992) KPSS tests for quotes and market depth. While the quote series are obviously integrated, we obtain conflictive findings for the depth series. The ADF tests reject the null hypothesis of a unit root in first level depth in 83% of all cases (across stocks and days), whereas the KPSS tests reject the stationarity in 67% of all cases. For higher level depth, the evidence against stationarity in depth is even higher. As discussed in Section 3.1, we explain these findings by the fact that order book depth is an inventory variable which over short horizons is strongly autocorrelated and tend to behave like an $I(1)$ process. On the other hand, aggressive trading and limit order arrivals create fluctuations in depth which are less predictable and reduce the strong persistence over longer intervals. Extreme changes arise, for instance, whenever first level depth is absorbed by an incoming order or, alternatively, is undercut by an incoming aggressive limit order, and thus the entire order book is shifted. Hence, from this discussion and the empirical findings we can conclude that depth might naturally contain stationary and non-stationary components where the latter tend to dominate over very short horizons. Given these results, it is in any case recommended to model depth as a non-stationary variable within a cointegrated VAR framework. As discussed in Section 3.1, this proceeding ensures consistency of parameter estimates even if depth might be stationary and, e.g., is fractionally cointegrated (see Johansen and Nielsen, 2010).

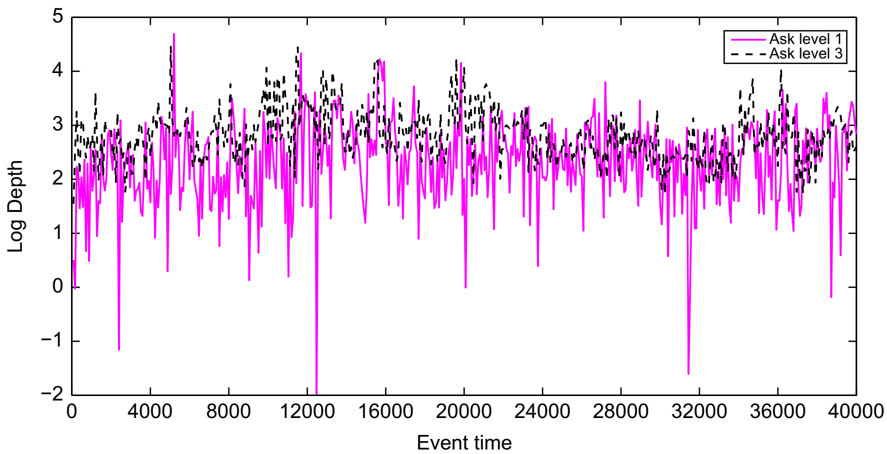

Fig. 5. Time series plot of log market depths (measured in thousand share units). Trading of Fortis, Euronext, Amsterdam, August 1, 2008.

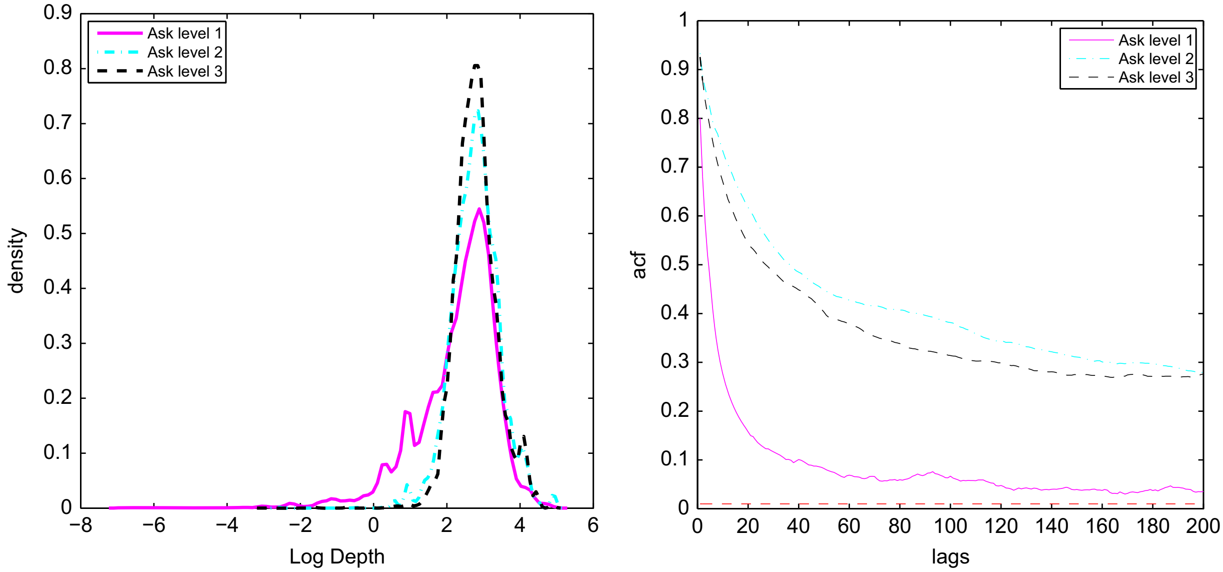

Fig. 6. *Left*: Kernel density estimates of (log) market depths. *Right*: Autocorrelation functions of (log) market depths. Trading of Fortis, Euronext, Amsterdam, August 1, 2008.

Table 4
Stationarity tests on quotes and market depths. Augmented Dickey–Fuller (ADF) tests and KPSS tests for the 30 selected stocks on each of the 43 trading days, i.e., 1290 time series for each variable. The chosen lag length is 50. The reported numbers are the sum of rejections at the 1%-level. In the ADF test, the null hypothesis is that there is an unit root in the process. In the KPSS test, the null hypothesis is that there is *no* unit root in the process.

Variables	p^a	p^b	$v^{a,1}$	$v^{a,2}$	$v^{a,3}$	$v^{b,1}$	$v^{b,2}$	$v^{b,3}$
ADF	8	4	1072	975	933	1087	975	949
(%)	(0.62)	(0.31)	(83.1)	(75.58)	(72.32)	(84.26)	(75.58)	(73.56)
KPSS	1284	1283	871	905	982	846	896	979
(%)	(99.53)	(99.45)	(67.51)	(70.15)	(76.12)	(65.58)	(69.45)	(75.89)

Table 5
Representative estimates of cointegrating vectors. The vectors are sorted according to their corresponding eigenvalues in Johansen's ML approach. The first two vectors are fixed to $\beta_1 = [0,\ldots,0,1,0]$ and $\beta_2 = [0,\ldots,0,0,1]$ representing stationary processes of trading indicators. Correspondingly, all entries in $\widehat{\beta}_3$ to $\widehat{\beta}_9$ associated with the trading indicator variables, *BUY* and *SELL*, are set to zero and are omitted. Trading of Fortis at Euronext, Amsterdam on August 1, 2008.

Variable	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$	$\hat{\beta}_7$	$\hat{\beta}_8$	$\hat{\beta}_9$
p^a	−0.9987	−1.0000	1.0000	−0.9989	1.0000	−1.0000	0.9399
p^b	1.0000	0.9853	−0.9968	1.0000	−0.9767	0.7048	−1.0000
$v^{a,1}$	−0.0173	0.1629	−0.0416	0.0222	−0.0803	0.0919	−0.1039
$v^{a,2}$	0.0070	−0.0486	0.0322	−0.0839	−0.1915	0.5869	−0.6605
$v^{a,3}$	−0.0070	0.0140	−0.0212	0.0108	0.2980	0.6104	−0.5413
$v^{b,1}$	−0.0081	−0.1412	−0.0398	0.0827	−0.0442	−0.0807	−0.0933
$v^{b,2}$	0.0003	0.0527	0.0430	0.2321	0.0167	−0.8162	−0.4652
$v^{b,3}$	−0.0002	−0.0342	−0.0212	−0.2988	0.0796	−0.9414	−0.3337

4.2. Estimated cointegration relations

For the sake of brevity, we refrain from showing the individual estimates of **A** and **B**. Nevertheless, it is interesting to highlight the estimated cointegration relations. According to Johansen's trace statistics we identify seven cointegration relations among quotes and depths. Table 5 shows the estimated cointegrating vectors for a representative trading day, where we omit the two known cointegrating vectors associated with the (stationary) trading indicators. Likewise we also omit the corresponding entries in the remaining cointegrating vectors as they are zero by construction. The resulting vectors are ordered according to their corresponding eigenvalues reflecting their likelihood contributions. Table 6 shows

Table 6
Representative estimates of the loading matrix. Values in parentheses are t-statistics. Trading of Fortis at Euronext, Amsterdam on August 1, 2008.

	$\hat{\beta}_3{}' y_{t-1}$	$\hat{\beta}_4{}' y_{t-1}$	$\hat{\beta}_5{}' y_{t-1}$	$\hat{\beta}_6{}' y_{t-1}$	$\hat{\beta}_7{}' y_{t-1}$	$\hat{\beta}_8{}' y_{t-1}$	$\hat{\beta}_7{}' y_{t-1}$
p_t^a	0.0818	−0.0104	−0.0084	0.0042	0.0022	−0.0004	0.0002
	(18.45)	(−2.34)	(−50.11)	(18.42)	(8.50)	(−2.65)	(0.89)
p_t^b	−0.0691	−0.0133	0.0026	0.0041	0.0007	−0.0004	−0.0000
	(−15.91)	(−3.07)	(15.58)	(18.33)	(2.77)	(−2.59)	(−0.07)
$v_t^{a,1}$	2.1666	−0.4160	0.5319	0.0108	0.1228	0.0013	0.0052
	(18.25)	(−3.50)	(118.70)	(1.76)	(17.52)	(0.31)	(0.75)
$v_t^{a,2}$	−0.2935	0.0512	−0.1776	0.0423	0.0995	−0.0074	0.0114
	(−7.44)	(1.29)	(−119.21)	(20.65)	(42.68)	(−5.52)	(5.01)
$v_t^{a,3}$	0.0589	−0.0117	0.1293	−0.0176	−0.1314	−0.0072	0.0079
	(1.58)	(−0.31)	(92.30)	(−9.14)	(−59.94)	(−5.72)	(3.71)
$v_t^{b,1}$	1.5157	0.4704	0.7016	−0.1435	0.0596	0.0017	0.0025
	(12.75)	(3.96)	(156.39)	(−23.26)	(8.49)	(0.42)	(0.37)
$v_t^{b,2}$	−0.2284	−0.0607	−0.2373	−0.1053	−0.0036	0.0087	0.0074
	(−5.99)	(−1.59)	(−165.07)	(−53.26)	(−1.58)	(6.72)	(3.40)
$v_t^{b,3}$	−0.0492	0.0060	0.1322	0.1253	−0.0176	0.0119	0.0034
	(−1.39)	(0.17)	(99.52)	(68.59)	(−8.44)	(10.01)	(1.66)

the estimated loading matrix, $\hat{\alpha}$, and corresponding t-statistics. We observe that not only quotes but also depth variables have a significant loading on most of the six cointegration relations.

Fig. 7 depicts the time series of the estimated cointegration relations. The series are quite different from that of the bid–ask spread (i.e., the difference between ask and bid quotes) which would be expected if depth does not belong to the cointegration vector and is also depicted in the figure. Compared to the spread which reflects a very discrete behavior, the cointegration relations are much more smooth. Nevertheless, as *any* linear combination of these vectors results into a further cointegration relation, it is required to formally test whether the estimated cointegration relations are indeed different from the bid–ask spread. The corresponding likelihood ratio test of the null hypothesis $R'\beta = 0$ with $R = [1, 1, 0, \ldots, 0]'$ rejects at a 1% significance level for all regressions (except one) for Fortis. Hence, we obtain significant evidence for depth being part of the cointegration relations influencing long-term equilibria of quotes and depth.[10]

Interpreting the estimated cointegrating vectors, we can derive several implications. The first five cointegration relations are mostly linear combinations of spreads and depths. Specifically, the first one is quite similar to the bid–ask spread as the coefficients for the depth variables are comparably small. The second cointegration relation seems to involve the balance of at-the-market depth since the coefficients of $v^{a,1}$ and $v^{b,1}$ are similar in magnitude and opposite in sign. The most interesting relationships are implied by the last two cointegrating vectors revealing relatively large (and different) coefficients associated with depth. This indicates that depth has a significant impact on the long-term relationship between quotes. Intuitively, the connection between ask and bid quotes becomes weaker (and thus deviates from the spread) if the depth is less balanced between both sides of the market. Hence, depth has a significant impact on quote dynamics and should be explicitly taken into account in a model for quotes. These findings support the idea of a cointegration model for *both* quotes and depth.

5. Estimated market impact

5.1. Limit orders placed at or behind the market

Consider the impact of an incoming at-the-market limit order as described in Scenario 1 in Section 3.2. Fig. 8 shows the impulse responses induced by buy and sell limit orders with a size equal to half of the depth at the best quotes.[11] The impulse response function starts at zero since such a limit order does not directly change the ask and bid. As expected, both ask and bid tend to significantly increase (decrease) after the arrival of a buy (sell) limit order. Induced by the cointegration setting, quotes converge to a (new) permanent level at which the information content of the incoming limit order is completely incorporated. The confidence intervals reflect that the shift is statistically highly significant.

We observe that quotes adjust relatively quickly reaching the new level after approximately 20 lags. Recall that time is measured in terms of limit order book activities. Hence, the adjustment speed measured in physical time ultimately depends on the underlying frequency of order activities and differs across the market. However, the fact that the speed of stock-specific quote adjustments (in terms of a ‘limit order clock’) is widely stable across the market, indicates that such a business time scale is appropriate for market-wide comparisons across stocks.

[10] It is well known that likelihood ratio tests on cointegration vectors tend to be biased towards rejecting the null hypothesis too often in finite samples, see, e.g., Gredenhoff and Jacobson (2001) and Haug (2002). However, given the high number of observations used in our study, these effects should not be too strong.

[11] In all figures illustrating impulse responses, the legend ‘A→B’ is interpreted to reflect ‘the impact on B induced by A’.

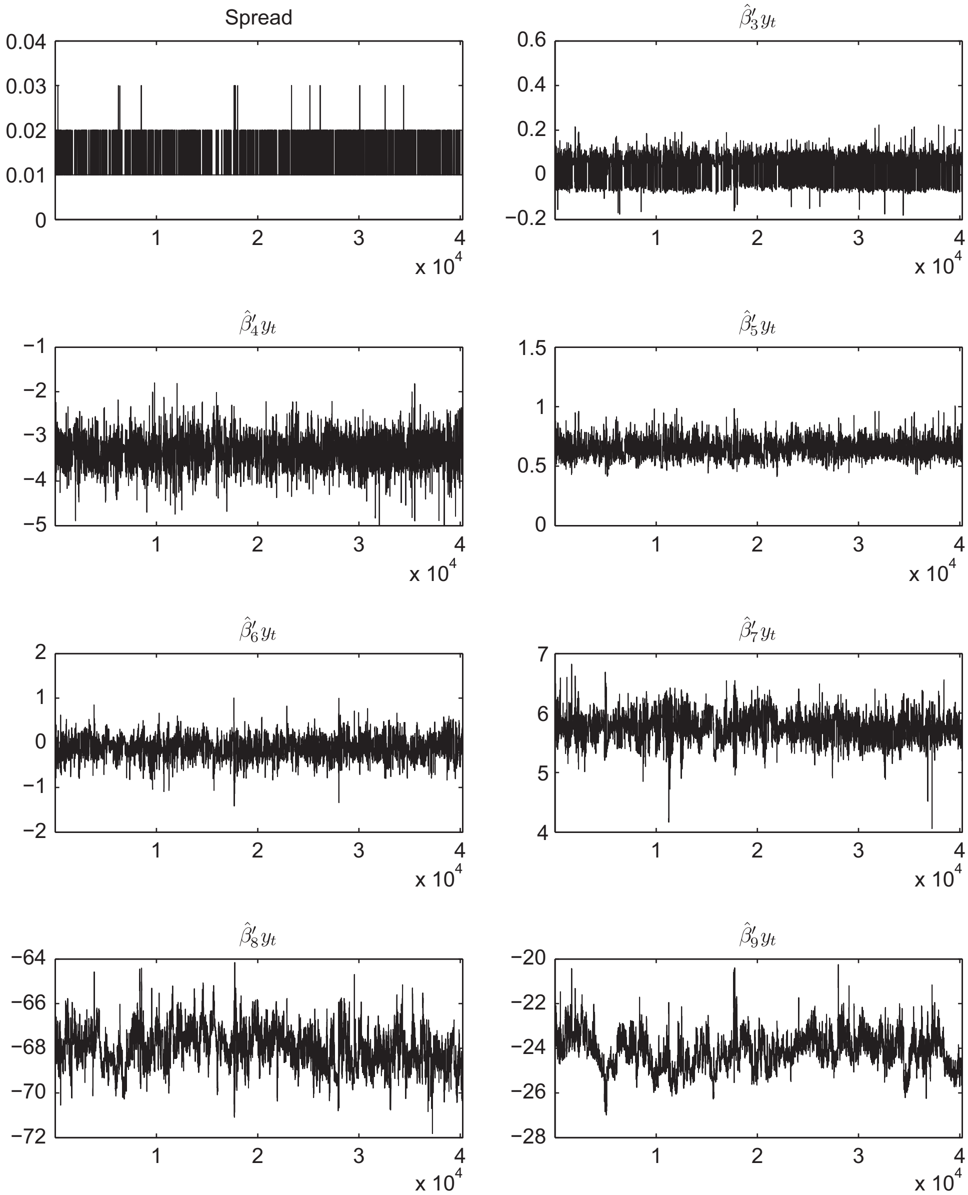

Fig. 7. Time series of estimated cointegration relations. The corresponding cointegrating vectors are documented in Table 5. We suppress the two cointegrating relationships associated with the trading indicator series. Trading of Fortis at Euronext, Amsterdam, on August 1, 2008.

An interesting fact is that after the arrival of a buy limit order, the bid tends to increase more quickly than the ask. A reverse effect is observed after the arrival of a sell limit order. This asymmetry introduces a one-sided and temporary decrease of the bid–ask spread. We explain this phenomenon by the fact that traders observing an incoming limit order on the same side of the market tend to compete for liquidity provision by undercutting quotes. Moreover, the higher depth at the bid generates a (delayed) liquidity demand on the ask side shifting upward the ask as well. We thus refer this phenomenon to be a liquidity-motivated effect.

Our findings can be interpreted in terms of pure market mechanisms. The market equilibrium is perturbed by a limit order in two ways. On one hand, the limit order indicates an investor's willingness to buy or sell and thus increases the supply or demand of the underlying asset. The market price changes to incorporate this temporary imbalance of supply and demand. On the other hand, an incoming limit order increases the supply of liquidity in the market. A narrowing of the spread reduces transaction costs and causes a re-balancing of supply and demand of liquidity. See, e.g., the simulation study by Yamamoto (2011) on the effects of the state of the limit order book on investors' strategies.

The significant permanent impact induced by an incoming limit order indicates that it contributes to price discovery. Thus, market participants perceive that limit orders carry private information which is in contrast to the common

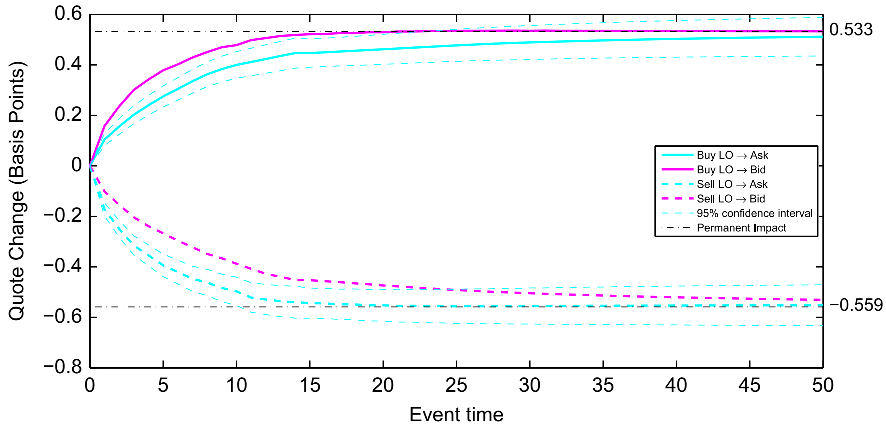

Fig. 8. Changes of ask and bid quotes induced by buy/sell limit orders placed at the market (level one) with a size equal to the half of the depth on the first level. The marked number on the vertical axes indicates the magnitude of the permanent impact. The blue dotted lines indicate the corresponding 95%-confidence intervals. Trading of Fortis at Euronext, Amsterdam in August 2008. LO: limit order. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

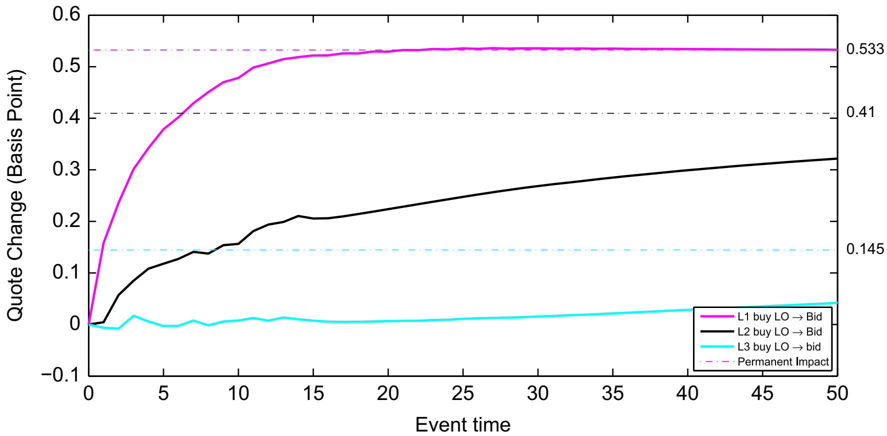

Fig. 9. Changes of bid quotes induced by buy limit orders placed at the market (level one) and behind the market (level two and three). The order size equals to half of that at the best bid. The initial order book equals the corresponding monthly average shown in Table 1. The marked number on the vertical axes indicates the magnitude of the permanent impact. Trading of Fortis at Euronext, Amsterdam in August 2008. L1: level one. L2: level two. L3: level three. LO: limit order.

assumption in theoretical literature that informed traders only take liquidity but do not provide it. On the other hand, it is supported by the experiment by Bloomfield et al. (2005) showing that informed traders use order strategies involving both market orders and limit orders to optimally capitalize their informational advantage and in line with Mike and Farmer (2008) and Chiarella et al. (2009) suggesting that there is a link between the properties of order flow and those of prices.

Given the setting of the book, we observe that a limit order increasing first level depth by 50% shifts quotes by 0.5 to 0.6 basis points. Though this effect is generally rather small, it is economically significant if the tick size is small. Moreover, note that the magnitude of the market impact is log-linear in the order size. In practice, a big limit order posted on a thin order book might affect the market much more strongly than in our scenarios.

In order to explore the role of the order's position in the book, Fig. 9 depicts the bid prices' reactions induced by buy limit orders placed at the market (level one) and behind the market (levels two and three). We observe a negative correlation between the magnitude of quote reactions and the orders' distance from the spread. The at-the-market limit order induces significantly faster market reactions than the behind-the-market limit order. Nonetheless, the long-term impact of level one and level two limit orders is only approximately 20% smaller. Hence, it turns out that behind-the-market orders can significantly shift the market though the quote adjustment is slower.[12] This result holds for level-two orders and (to a weaker extent) for level-three orders. However, for orders posted deeper in the book virtually no market impacts can be identified.

[12] In order to improve the graphical illustrations, we refrain from showing the corresponding confidence intervals. They are quite similar to those shown in Fig. 8.

Eom et al. (2009) find evidence that traders could have made extra profits using microstructure-based manipulations on the Korean Exchange (KRX) during a period between 2001 and 2002. In this period, KRX disclosed the total quantity on each side of the order book without fully disclosing the prices at which these orders have been placed. The manipulation strategy resulted in placing huge numbers of behind-the-market limit orders on the opposite side of the market inducing price moves in the favorite direction without having these orders executed. Our finding shows that this kind of manipulation is indeed possible. However, whether it is economically profitable in Euronext trading, ultimately depends on (relative) order sizes. In order to move quotes in her favorite direction, the trader has to submit rather big limit orders close to the market. Then, she faces the risk that these orders may be picked up.

5.2. Limit orders placed inside of the spread

Limit orders placed inside of the bid–ask spread perturb the LOB dynamics in a more complex way. Apart from providing liquidity to the order book, they directly improve the ask or bid. This quote adjustment induces a reduction of the spread, establishes a new best quote level and correspondingly shifts all depth levels on the corresponding side of the book upward (or downward, respectively). The system seeks the new equilibrium on a path recovering from the immediate quote change and simultaneously re-balancing liquidity. Given our setting, we assume that a buy limit order inside of the spread induces a 5 basis points increase of the bid. However, as shown in the left plot of Fig. 10, the long-run price impact is just 1.8 basis points. The immediate quote movement is reverted back by approximately 65%. This is induced either by sell trades picking up the posted volume or by cancellations on the bid side. Similarly, liquidity demand on the ask side shifts the ask upward by 1.8 basis points. Hence, overall we observe an asymmetric re-balancing of quotes and a corresponding re-widening of the spread.

The right plot of Fig. 10 compares the effects of buy limit orders of different sizes but with same limit price posted inside of the bid–ask spread and thus improving bid quotes again by 5 basis points. We observe quite different impulse response patterns in dependence of the order size. In case of a comparably small order, the posted volume is quickly picked up, shifting back the bid quote. In contrast, large volumes overbidding the prevailing quote cause a long-term upward movement of the bid. Relative to the initial shift of the bid we observe a further approximately 20% price increase. Hence, extraordinary large orders are not likely to be picked up and induce a rather strong buy pressure moving the market upwards. For smaller (though still comparably large) orders, adverse selection and signalling effects seem to counter-balance each other. As a consequence, the bid quote is hardly changed and the long run effect is close to the immediate price improvement.

5.3. Market impact of trades

Fig. 11 shows the market impacts induced by buy and sell market orders. We assume that the trade sizes correspond to 50% of the prevailing depth. Consequently, these market orders do not ‘walk through’ the limit order book and thus ask and bid quotes are unaffected. The quote adjustments shown in Fig. 11 are subsequent quote responses to trade arrivals. Both bid and ask increase (decrease) sharply after the arrival of a buy (sell) market order. Hence, the arrival of a buy (sell) market order induces aggressive posting on the bid (ask) side resulting in further buy (sell) market orders and limit orders posted inside of the spread. Similar to the findings for limit orders, we find evidence for asymmetric adjustments of the two sides of the market. It turns out that buy market orders shift ask quotes more quickly and strongly than bid quotes. The reverse is true for sell market orders. This result indicates that trades temporarily increase spreads which is in contrast to the effects induced by limit orders. Engle and Patton (2004) report similar findings by analyzing quote data from the NYSE. They show that trades have a positive impact on spreads. As they directly impose the spread as the underlying cointegration relation of quotes, the price impact of trades on spreads is transitory in their model. Using impulse-response analysis based on a structural VEC model, Escribano and Pascual (2006) also find that spreads permanently widen after the arrival of trades. Note that these effects contradict implications of asymmetric-information-based market microstructure models, such as Glosten and Milgrom (1985) and Easley and O’Hara (1992), where trades should resolve the uncertainty regarding existing information and should result in declining spreads.

The left plot of Fig. 12 depicts quote reactions induced by an aggressive market order ‘walking through’ the book (Scenario 4 in Section 3.2). It absorbs the best ask level and shifts the best quote to the originally second best level which is assumed to be 10 basis points higher than the previous best ask. Similar to the effects induced by aggressive limit orders we observe that the initial shift of the best ask is reverted back by approximately 35% inducing a long-run ask increase of 6.5 basis points. Simultaneously, aggressive posting on the bid side shifts bid quotes upward. Hence, the initially widened spread reverts back in an asymmetric way causing more quote movements on the bid side than on the ask side. The responses mirror the corresponding effects induced by aggressive buy limit orders (see Fig. 10), where the spread is initially narrowed and then asymmetrically re-widened causing also more movements on the bid side than on the ask side.

The right plot of Fig. 12 compares the market impacts on ask quotes induced by a buy market order in situations of different depths behind the market. It is assumed that the order just absorbs the first ask level and thus induces an instantaneous ask price increase by 5 basis points. In line with the results discussed above, in all three scenarios the initially shifted ask quote is reverted back. However, it turns out that the magnitude of this quote reversion critically depends on the prevailing depth behind the market. In fact, the existence of a huge level-two-depth reverts the ask quote

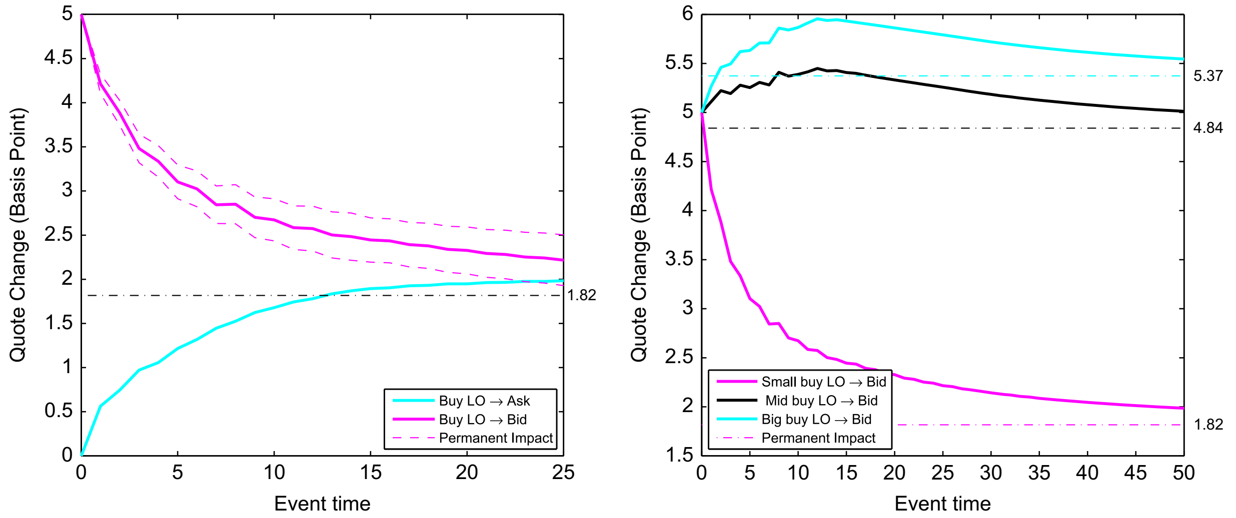

Fig. 10. *Left*: changes of quotes induced by buy limit orders placed inside of the spread with a size equal to the depth at the bid. *Right*: changes of bid induced by buy limit orders placed inside of the spread with different sizes. The initial order book equals the corresponding monthly average shown in Table 1. Small size: depth at the bid. Mid size: 7 times of the depth at the bid. Big size: 15 times of the depth at the bid. Trading of Fortis at Euronext, Amsterdam in August 2008. LO: limit order.

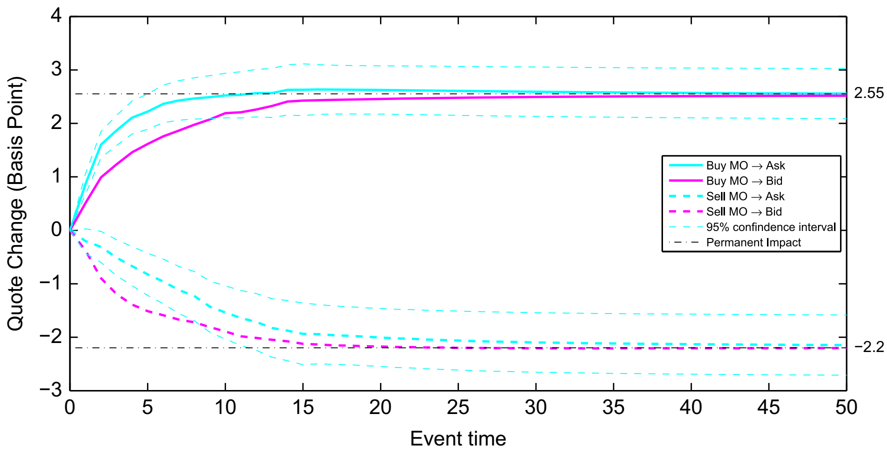

Fig. 11. Changes of quotes induced by buy/sell market orders (buyer-/seller-initiated trades) with a size equal to half of the depth on their corresponding first levels. The marked number on the vertical axes indicates the magnitude of the permanent impact. Trading of Fortis at Euronext, Amsterdam in August 2008. MO: Market order.

back by approximately 60%. We explain this fact by a strong sell pressure induced by huge sell volume queued on the ask side. Conversely, in case of only little depth prevailing behind the market, the existing sell pressure is weaker causing the incoming buy order to (upward) shift the market more strongly. In the extreme case of a very thin market, we even observe an additional quote increase.

A practical problem faced by many market participants is the fundamental choice between posting a market order and a limit order. A direct comparison of the market impacts induced by these two types of orders is shown in Fig. 13. In both cases, the posted order does not *directly* change quotes. We observe that the resulting long-run effect of trades is significantly greater than that of an equal-size limit order. Actually, the price shift induced by a market order is approximately four times larger than that of a comparable limit order. This finding is consistent with the theoretical prediction by Rosu (2010). Moreover, market orders also cause quicker market reactions. Finally, inferred from the 'gap' between ask and bid curves, market orders change the spread more dramatically than limit orders. Hence, the willingness to cross the bid–ask spread is a stronger signal for private information than that induced by a comparable limit order.

Note that the comparison holds for 'normal' order types placed on the best quote, but not necessarily for more aggressive orders. As discussed above, the long-term effects of aggressive limit orders and market orders critically depend on their (relative) size and the current state of the book. Therefore, an ultimate comparison of market impacts induced by

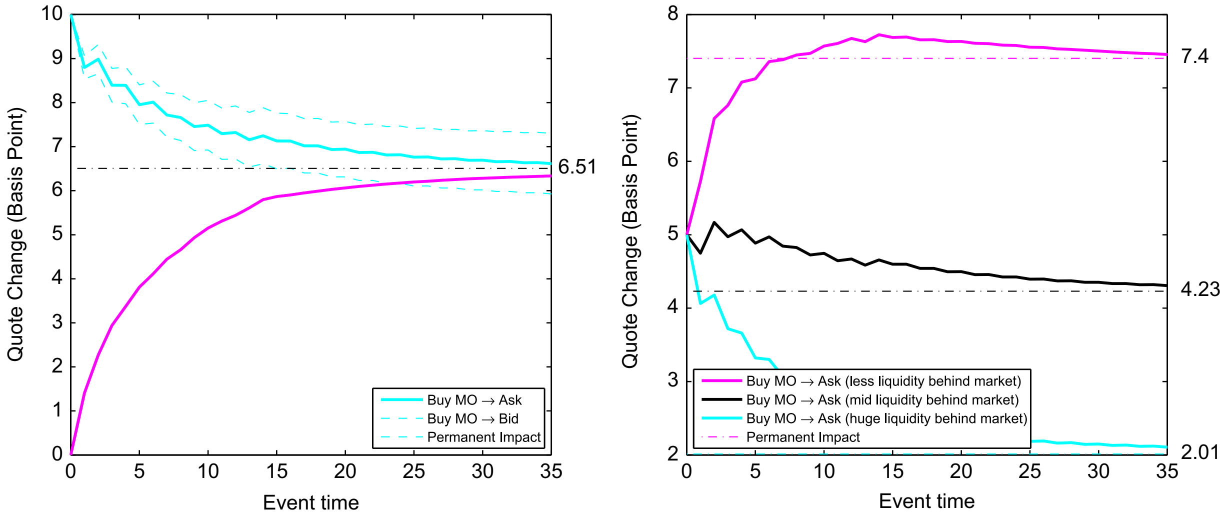

Fig. 12. *Left*: changes of bid and ask quotes induced by an aggressive buy market order with a size exceeding the depth at the best ask by 20%. The second best ask is assumed to be 5 basis points higher than the ask, where the depths behind the market are 1.5 times of the depth at the market. *Right*: changes of ask quotes induced by an aggressive buy market order with a size equal to the depth at the ask when there is different depth at the second best ask. Case 1: the depth at the second best ask is 10% of that at the ask; Case 2: the depth at the second best ask equals that at the ask; Case 3: the depth at the second best ask is 500% of that at the ask. The marked number on the vertical axes indicate the magnitude of the permanent impact. Trading of Fortis at Euronext, Amsterdam in August 2008. MO: market order.

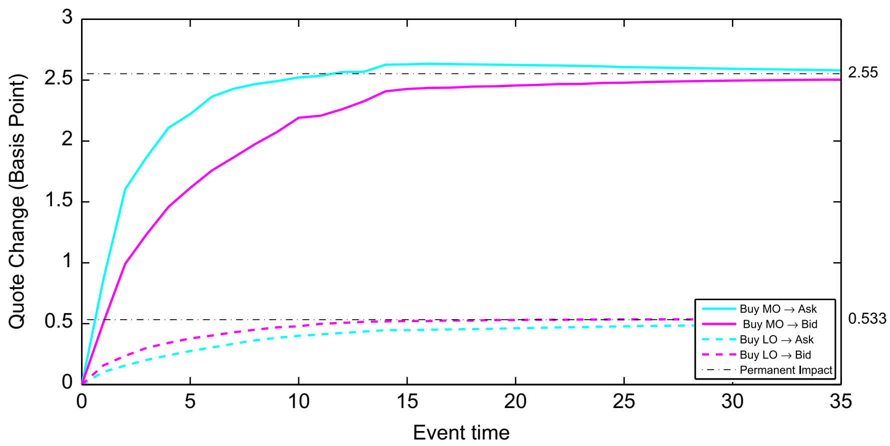

Fig. 13. Changes of ask and bid quotes induced by a buy market order and a buy limit order of similar size placed at the market. The order size is half of the depth at the best bid. The depths at the best bid and the best ask in the order book are assumed to be equal. Trading of Fortis at Euronext, Amsterdam in August 2008. LO: limit order; MO: market order.

both types of orders under comparable conditions is rather difficult. Nevertheless, our results show that limit orders do have a significant long-term effect and can significantly 'scare' the market.

5.4. *Robustness of results*

Selecting the appropriate lag order in VAR models is cumbersome in practice when a substantial cross-section of stocks is analyzed over a comparably long period. In order to analyze the sensitivity of our results regarding the choice of the lag order in the VAR model, Fig. 14 compares the market impacts of a buy limit order and that of a normal buy market order predicted by a VAR(15) model with those induced by a VAR(10) specification using trading of Fortis in August 2008. It turns out that despite a misspecification of the lag length and remaining serial correlation in the residuals, the impulse response estimates of a VAR(10) are quite close to that of a VAR(15). This is in line with results reported by Jorda (2005) using a VAR(2) to estimate impulse-response functions of an underlying VAR(12) model.

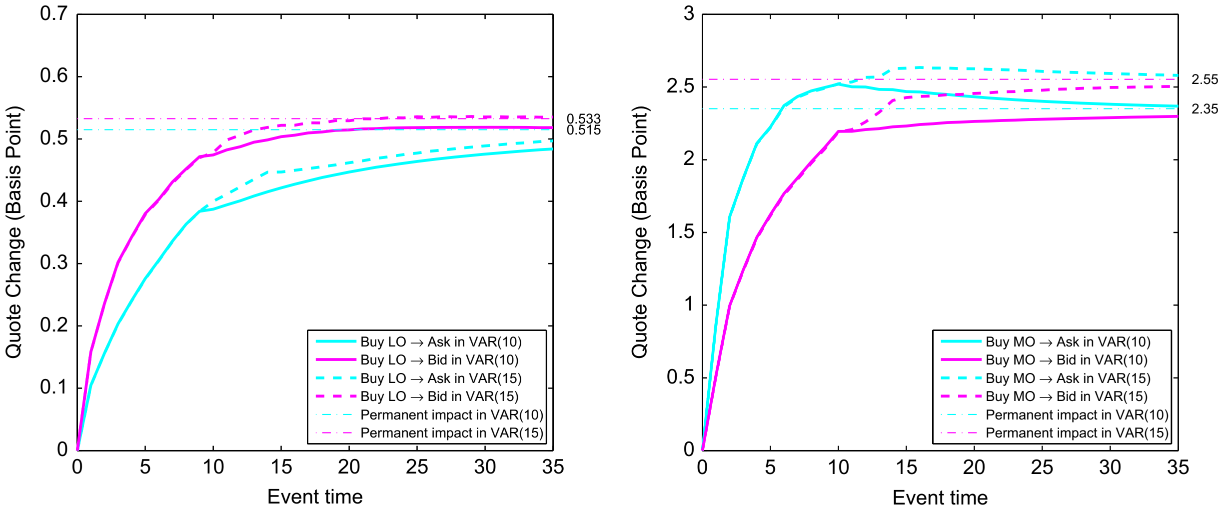

Fig. 14. Robustness of results. Market impacts of a bid limit order estimated by a VAR(15) and a VAR(10) specification. Trading of Fortis, Euronext Amsterdam in August 2008.

5.5. *Cross-sectional evidence*

The complete empirical analysis has been conducted for 29 other stocks traded at Euronext Amsterdam using a VARX(15) specification. The corresponding results are shown in the Appendix on the companion web site at http://amor.cms.hu-berlin.de/~huangrui/project/impact_of_orders/. It turns out that the results reported in the previous sections are qualitatively stable and representative for a wide cross-section of stocks. Nevertheless, we observe that the *magnitudes* of market impacts vary across the market and seem to be driven by underlying liquidity characteristics. To gain insights into these relationships, we run a simple cross-sectional regression of absolute average market impacts on the average stock-specific trading frequency, trading volume as well as the minimum tick size. I.e.,

$$M_i = \gamma_0 + \gamma_1 N_i + \gamma_2 S_i + \gamma_3 V_i + \varepsilon_i, \tag{13}$$

where M_i denotes the absolute permanent impact of stock i induced by a buy/sell order, N_i is the average number of trades per day, S_i represents the normalized tick size, and V_i denotes the normalized trade volume per day. Particularly,

$$S = \frac{\text{tick size} \times 100}{\text{average of closing prices}}, \quad V = \frac{\text{adjusted trading volume per day}}{\text{number of outstanding shares}} \times 100.$$

The scenarios we consider below are similar to those studied in Section 3.2. The initial order book for each stock equals its monthly average.

Scenario 'normal limit order' and 'normal market order': These scenarios are identical to that in Section 3.2.

Scenario 'aggressive limit order': An incoming order of a size which is half to the depth at the corresponding best price is posted inside of the spread and improves the corresponding quote by one tick.

Scenario 'aggressive market order': An incoming market order with a size equal to the depth at the corresponding best price and thus absorbing the first level in the book.

For every scenario, we consider average market impacts of both buy and sell orders for 30 stocks estimated over two months resulting in 120 observations for each regression.

Table 7 reports the corresponding estimation results for two versions of the model: one with included trading volume and one without. The high R^2 values, ranging between 67% and 80%, show that most of the cross-sectional variation of the market impact can indeed be explained by the three explanatory variables. It turns out that the trading volume (though its parameter is significant) does not provide much explanatory power. This result indicates that the trading frequency rather than the trade size drives the strength of market responses to limit order arrivals. Furthermore, we observe that the trading frequency has a negative influence on the market impact of limit orders. Hence, in case of a slower trading, a single order conveys more information.

The tick size is positively related to the magnitude of permanent impacts in all scenarios. For aggressive limit orders, this finding is not surprising as the implied price improvement is (relatively) higher for stocks trading on larger tick sizes. Since in these cases, also the spreads between best and second best quotes are higher, the immediate price shift by the arrival of an aggressive market order is larger as well. In the scenarios 'normal limit order' and 'normal market order', a higher tick size and thus an increase of the price discreteness makes it more likely that investors are forced to under-react or over-react in response to incoming information inducing higher deviations between quoted prices and the 'true'

Table 7
Cross-sectional analysis of market impacts over the market. We consider market impacts of both buy and sell orders for 30 stocks estimated over August and September 2008. The regression is $M_i = \gamma_0 + \gamma_1 N_i + \gamma_2 S_i + \gamma_3 V_i + \varepsilon_i$, where M_i denotes the absolute permanent impact of stock i induced by a buy/sell limit order, N_i is the average number of trades per day, S_i represents the normalized tick size, and V_i denotes the normalized trade volume per day. The numbers in brackets denote heteroskedasticity robust t-statistics according to White (1980).

Scenario	γ_0	γ_1	γ_2	γ_3	R^2
'Normal limit order'	0.0033	−0.0013	0.0418	–	0.69
	(19.02)	(−7.30)	(33.72)		
	0.0027	−0.0015	0.040	0.0011	0.73
	(13.77)	(−11.35)	(34.85)	(6.5)	
'Aggressive limit order'	0.005	−0.0017	0.096	–	0.79
	(13.12)	(−5.93)	(23.04)		
	0.004	−0.002	0.095	0.001	0.80
	(11.98)	(−7.92)	(22.80)	(8.34)	
'Normal market order'	0.037	−0.016	0.17	–	0.57
	(21.94)	(−9.26)	(6.33)		
	0.030	−0.018	0.151	0.013	0.67
	(14.59)	(−16.74)	(5.56)	(7.49)	
'Aggressive market order'	0.049	−0.020	0.474	–	0.68
	(19.89)	(−8.24)	(14.46)		
	0.039	−0.023	0.449	0.018	0.75
	(13.03)	(−15.06)	(14.07)	(7.22)	

underlying efficient price. Our findings show that in these situations, investors rather tend to over-react after the arrival of an order.

6. Conclusion

In this paper, we quantify the market impact of incoming limit orders in a limit order book market. Best bid and ask quotes as well as three levels of depth on both sides of the market are modelled based on a cointegrated VAR system. Incoming limit orders are represented in terms of shocks to the system. Limit order characteristics as well as the corresponding state of the book are captured by the specific design of the shock vector. This allows us to distinguish between limit orders of different aggressiveness (reflected by their distance to the market) and different sizes as well as between different states of the book. The market impacts on ask and bid prices are quantified by the estimated impulse response function using appropriate statistical inference.

Employing this framework we analyze the limit order book processes of 30 stocks traded on Euronext Amsterdam over two months in 2008. The model is estimated using the highest possible frequency accounting for all order book changes during continuous trading. Parameter estimates and diagnostics indicate that the proposed model captures the high-frequency order book dynamics quite well.

Based on the empirical analysis we can summarize the following findings: First, we find clear evidence for cointegration relations between ask and bid quotes and corresponding depths. While some cointegration relations are similar to the bid–ask spread, others show that depth has a distinct effect on quote dynamics and on the connection between ask and bid quotes. Second, limit orders do have significant long-term effects on quotes. This is even true for limit orders placed behind the market though these effects decline with the limit order's distance to the market. While incoming limit orders temporarily decrease the spread, market orders induce a temporary widening. Third, the speed of spread convergence as well as the direction of price movements after the arrival of aggressive limit orders undercutting (or overbidding, respectively) best ask and bid prices strongly depends on the incoming limit order's size. While small orders seem to face adverse selection risks and are likely to be picked up quickly, for larger orders information signalling effects seem to dominate pushing the market in the opposite direction. Fourth, the decrease (increase) of spreads after the arrival of an aggressive limit (market) order is reverted back asymmetrically inducing more quote movements on the side where the order has been placed. Fifth, the long-run market impact of aggressive market orders walking through the book rises with the queued depth behind the market. Sixth, the effects are qualitatively remarkably stable over the cross-section of the market. Variations in the magnitudes of market impacts are well explained by the underlying stock-specific trading frequency and minimum tick size.

Our empirical results also show that the proposed framework is useful and appropriate to capture order book dynamics on high frequencies. By modelling quotes *and* several levels of depth, the model implicitly captures also the multivariate dynamics of mid-quotes, returns, spreads, spread changes as well as depth imbalances. In this sense, the suggested high-frequency cointegrated VAR model can serve as a workhorse for various applications in this area.

Acknowledgments

For helpful comments and discussions we thank Jean-Philippe Bouchaud, Carl Chiarella, David Easley, Robert Engle, Maureen O'Hara, Joel Hasbrouck, Albert Menkveld, Roel Oomen, Giorgio Valente, two anonymous referees and the participants of the Conference on ‘‘Individual Decision Making, High Frequency Econometrics and Limit Order Book Dynamics’’ in Warwick, 2009, the Market Microstructure Conference at the University of Technology, Sydney, 2010, the Conference ‘‘Market Microstructure, Confronting Many Viewpoints’’, in Paris, 2010, and seminars at the University of Leicester, Humboldt-Universität zu Berlin and Capital Fund Management, Paris. This research is supported by the Deutsche Bank AG via the Quantitative Products Laboratory and the Deutsche Forschungsgemeinschaft (DFG) via the Collaborative Research Center 649 ‘Economic Risk’.

Appendix A. Adaptive time window matching algorithm

In our database, trade data and order book data are recorded in separate files stemming from different recording systems. As a result, the time stamps in the two data sets have different time distances to exchange time. In accordance with the institutional settings of Euronext, we design an adaptive time window matching algorithm which contains three main steps.

Step 1: Exact matching. The algorithm picks up a time stamp of a trade and opens a specified time window, e.g., $[-10, 10]$ s around this time stamp. Then, a procedure picks every order book record in this time window and performs the following analysis: If (i) the trade price equals to the best bid (ask) price and the difference of the best bid (ask) size between this order book record and the previous one equals the trade size or (ii) the trade price equals the previous best bid (ask) price, the best bid (ask) size equals the trade size and the best bid (ask) price decreases (increases), it matches this order book record with the corresponding trade and records the delay time between the trade and the order book. If no match is achieved for all order book records in the time window, the trade remains to be unmatched.

Step 2: Inexact matching. The algorithm picks up an unmatched trade record's time stamp and opens a time window of size which is twice the average delay time computed in Step 1. If (i) the trade price equals the best bid (ask) price and the best bid (ask) size is less than the previous one or (ii) the best bid (ask) price decreases (increases), it matches the trade with the current order book. If no match is achieved for all order book records in the time window, the trade remains to be unmatched.

Step 3: Round time matching. The algorithm picks up an unmatched trade and matches it with an order book record that is closest to the trade's time stamp plus the average delay time.

Fig. A1 gives the histogram of the delay time between trades and their corresponding order book records. The delay time is computed in Step 1 by which 1609 (sub-)trades have been exactly matched with their corresponding order books inside a $[-5,5]$ s time windows. The average delay time is -0.185 s, i.e., trades are on average recorded 185 ms before the corresponding order book.

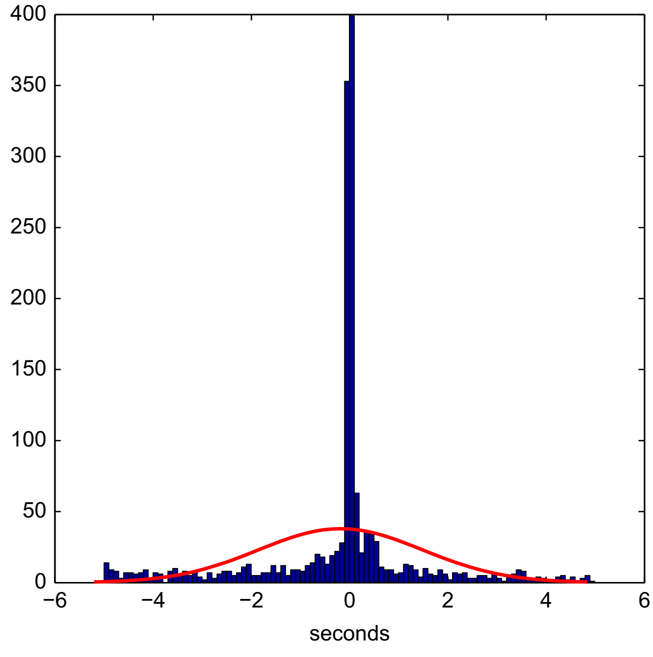

Fig. A1. Histogram of the delay time between the trade and its corresponding order book record. Trading of Fortis, Euronext Amsterdam on August 1, 2008.

Appendix B. FIML estimator for cointegrating vectors

Model (1) is estimated by the Full Information Maximum Likelihood (FIML) estimator proposed by Johansen (1991) and Johansen and Juselius (1990). Let $z_{0t} := \Delta y_t$, and $z_{1t} := y_{t-1}$. Further let z_{2t} be the vector of stacked variables,

$$z_{2t} := (\Delta y_{t-1}, \ldots, \Delta y_{t-p+1}, x_{t-1}, \ldots, x_{t-s}, 1)'$$

with corresponding parameter vector $\Gamma = (\Gamma_1, \ldots, \Gamma_{p-1}, \mu)$. Define the product moment matrices

$$M_{ij} := T^{-1} \sum_{t=1}^{T} z_{it} z_{jt}', \quad i,j = 0,1,2,$$

where T is the number of observations. Moreover, let

$$S_{ij} := M_{ij} - M_{i2} M_{22}^{-1} M_{2j}.$$

We then solve the generalized eigenvalue problem

$$|\lambda S_{11} - S_{10} S_{00}^{-1} S_{01}| = 0$$

for the eigenvalues $1 > \hat{\lambda}_1 > \cdots > \hat{\lambda}_K > 0$ and corresponding eigenvector $\hat{V} = (\hat{v}_1, \ldots, \hat{v}_K)$ which is normalized by $\hat{V} S_{11} \hat{V} = I_K$. Johansen's (1991) trace test or maximum eigenvalue test can be used to determine the underlying cointegration rank r. Under the hypothesis that there exist r cointegration relations, the $K \times r$ cointegrating matrix β is estimated by

$$\hat{\beta} = (\hat{v}_1, \ldots, \hat{v}_r)$$

with corresponding maximized log-likelihood function

$$l_{\max}(\hat{\beta}) = -\frac{T}{2}\left(\ln|S_{00}| + \sum_{i=1}^{r} \ln(1-\hat{\lambda}_i)\right). \tag{B.1}$$

The magnitude of $\hat{\lambda}_i$ can be interpreted as a measure of the stationarity of the product $\hat{\beta}_i' y_t$. The larger $\hat{\lambda}_i$, the closer the stochastic properties of the underlying relationship to that of a stationary process. The parameters α and Γ are estimated by OLS after inserting $\hat{\beta}$ into Eq. (1) and computing $\hat{\Sigma}_u$ as $\hat{\Sigma}_u = S_{00} - \hat{\alpha}\hat{\alpha}'$.

When some of the cointegration relations are known, we can partition the cointegrating matrix as

$$\beta = (b, \varphi),$$

where b contains known cointegrating vectors and φ contains the unknown ones. Denote

$$S_{ij.b} = S_{ij} - S_{i1}(b' S_{11} b)^{-1} S_{1j}.$$

Then, φ can be estimated similarly by solving the generalized eigenvalue problem

$$|\lambda b_\perp' S_{11.b} b_\perp - b_\perp' S_{10.b} S_{00.b}^{-1} S_{01.b} b_\perp| = 0.$$

Appendix C. Transform estimates from VECM to VAR

Following Lütkepohl and Reimers (1992), the parameters of Eq. (1) can be easily transformed to Eq. (2). In this context, we define a transformation matrix

$$D = \underbrace{\begin{bmatrix} I_K & 0 & 0 & & \cdots & 0 & 0 & 0 \\ I_K & -I_K & 0 & & \cdots & 0 & 0 & 0 \\ 0 & I_K & -I_K & \ddots & & 0 & 0 & 0 \\ \vdots & \vdots & & \ddots & \ddots & & \vdots & \vdots \\ 0 & 0 & 0 & & \ddots & 0 & 0 & 0 \\ 0 & 0 & 0 & & \cdots & -I_K & 0 & 0 \\ 0 & 0 & 0 & & \cdots & I_K & -I_K & 0 \\ & & 0 & & & & & 1 \end{bmatrix}}_{(KP+1)\times(KP+1)}$$

such that

$$[A_1,\ldots,A_p,\mu]=[\alpha\beta',\Gamma]D+J^*, \tag{C.1}$$

where $J^*:=[I_K:0:\cdots:0]$ is a $K\times(Kp+1)$ matrix. The theorem below provides a consistent estimator of A and B:

Theorem 1 (Lütkepohl and Reimers, 1992). *Let $\hat{\alpha}$, $\hat{\beta}$, $\hat{\Gamma}$ and $\hat{\Sigma}_u$ denote the FIML estimates of the parameters of model* (1). *Moreover, $\hat{A}_1,\ldots,\hat{A}_p,\hat{\mu}$ are computed by the transformation in* (C.1). *Then,*

$$\sqrt{T}[\mathrm{vec}(\hat{A}_1,\ldots,\hat{A}_p,\hat{\mu})-\mathrm{vec}(A_1,\ldots,A_p,\mu)]\xrightarrow{d}\mathcal{N}(0,\Sigma_{AB}), \tag{C.2}$$

where

$$\Sigma_{AB}=D'\begin{bmatrix}\beta & 0\\ 0 & I_{K(p-1)+1}\end{bmatrix}\Omega^{-1}\begin{bmatrix}\beta' & 0\\ 0 & I_{K(p-1)+1}\end{bmatrix}D\otimes\Sigma_u,$$

$$\Omega=\mathrm{plim}\frac{1}{T}\begin{bmatrix}\beta' M_{11}\beta & \beta' M_{12}\\ M_{21}\beta & M_{22}\end{bmatrix}$$

are consistently estimated by

$$\hat{\Sigma}_{AB}=D'\begin{bmatrix}\hat{\beta} & 0\\ 0 & I\end{bmatrix}\hat{\Omega}^{-1}\begin{bmatrix}\hat{\beta}' & 0\\ 0 & I\end{bmatrix}D\otimes\hat{\Sigma}_u,$$

$$\hat{\Omega}=\begin{bmatrix}\hat{\beta}' M_{11}\hat{\beta} & \hat{\beta}' M_{12}\\ M_{21}\hat{\beta} & M_{22}\end{bmatrix}.$$

References

Bloomfield, R., O'Hara, M., Saar, G., 2005. The "make or take" decision in an electronic market: evidence on the evolution of liquidity. Journal of Financial Economics 75, 165–199.
Boulatov, A., George, T., 2008. Securities trading when liquidity providers are informed. Working paper, C. T. Bauer College of Business, University of Houston.
Chiarella, C., Iori, G., Perelló, J., 2009. The impact of heterogeneous trading rules on the limit order book and order flows. Journal of Economic Dynamics & Control 33, 525–537.
Cont, R., Kukanov, A., Stoikov, S., March 2011. The Price Impact of Order Book Events. SSRN eLibrary.
Dufour, A., Engle, R.F., 2000. Time and the impact of a trade. Journal of Finance 55, 2467–2498.
Easley, D., O'Hara, M., 1992. Time and the process of security price adjustment. Journal of Finance 47 (2), 577–605.
Eisler, Z., Bouchaud, J., Kockelkoren, J. The price impact of order book events: market orders, limit orders and cancellations. Quantitative Finance, doi:10.1080/14697688.2010.528444, in press.
Engle, R., Patton, A., 2004. Impact of trades in an error-correction model of quote prices. Journal of Financial Markets 7, 1–25.
Eom, K., Lee, E., Park, K., 2009. Microstructure-Based Manipulation: Strategic Behavior and Performance of Spoofing Traders. SSRN eLibrary.
Escribano, A., Pascual, R., 2006. Asymmetries in bid and ask responses to innovation in the trading process. Empirical Economics 30, 913–946.
Glosten, L., Milgrom, P., 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. Journal of Financial Economics 14, 71–100.
Gredenhoff, M., Jacobson, T., 2001. Bootstrap testing linear restrictions on cointegrating vectors. Journal of Business and Economic Statistics 19 (1), 63–72.
Hansen, P.R., Lunde, A., 2006. Realized variance and market microstructure noise. Journal of Business and Economic Statistics 24 (2), 127–161.
Härdle, W.K., Hautsch, N., Mihoci, A., 2009. Modelling and forecasting liquidity supply using semiparametric factor dynamics. Discussion Paper 2009/18, Collaborative Research Center 649 "Economic Risk", Humboldt-Universität zu Berlin.
Hasbrouck, J., 1991. Measuring the information content of stock trades. Journal of Finance 46 (1), 179–207.
Haug, A.A., 2002. Testing linear restrictions on cointegrating vectors: sizes and powers of Wald and likelihood ratio tests in finite samples. Econometric Theory 18 (2), 505–524.
Hautsch, N., Huang, R., 2011. Limit order flow, market impact and optimal order sizes: evidence from Nasdaq total view-itch data. SFB 649 Discussion Paper 2011-056, Humboldt University, Berlin, Germany.
Johansen, S., 1991. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. Econometrica 59 (6), 1551–1580.
Johansen, S., 1995. Likelihood-Based Inference in Cointegrated Vector Autoregressive Models. Oxford University Press.
Johansen, S., Juselius, K., 1990. Maximum likelihood estimation and inference on cointegration—with applications to the demand for money. Oxford Bulletin of Economics and Statistics 52, 169–210.
Johansen, S., Nielsen, M., 2010. Likelihood inference for a fractionally cointegrated vector autoregressive model. Research Paper 2010-24, CTEATES.
Jorda, O., 2005. Estimation and inference of impulse response by local projections. American Economic Review 95, 161–182.
Kwiatkowski, D., Phillips, P., Schmidt, P., Shin, Y., 1992. Testing the null hypothesis of stationarity against the alternative of a unit root. Journal of Econometrics 54, 159–178.
Lee, C., Ready, M., 1991. Inferring trade direction from intraday data. Journal of Finance 46 (2), 733–746.
Ljung, G., Box, G., 1978. On a measure of lack of fit in time series models. Biometrika 65 (2), 297–303.
Lo, I., Sapp, S., 2006. A structural error-correction model of best prices and depths in the foreign exchange limit order market. Working Paper 06-8, Bank of Canada. ⟨http://EconPapers.repec.org/RePEc:bca:bocawp:06-8⟩.
Lütkepohl, H., 1990. Asymptotic distributions of impulse response functions and forecast error variance decompositions of vector autoregressive models. Review of Economics and Statistics 72, 116–125.
Lütkepohl, H., Reimers, H., 1992. Impulse response analysis of cointegrated systems. Journal of Economic Dynamics & Control 16, 53–78.
Mike, S., Farmer, J., 2008. An empirical behavioral model of liquidity and volatility. Journal of Economic Dynamics & Control 32, 200–234.
Parlour, C., Seppi, D., 2008. Limit order markets: a survey. In: Boot, A., Thakor, A. (Eds.), Handbook of Financial Intermediation & Banking, pp. 63–96.
Potters, M., Bouchaud, J., 2003. More statistical properties of order books and price impact. Physica A 324, 133–140.
Rosu, I., 2010. Liquidity and Information in Order Driven Markets. SSRN eLibrary.

Russell, J., Kim, T., 2010. A new model for limit order book dynamics. In: Bollerslev, T., Russell, J., Watson, M. (Eds.), Volatility and Time Series Econometrics, Essays in Honor of Robert Engle, Oxford University Press, pp. 354–364.

Said, E., Dickey, D., 1984. Testing for unit roots in autoregressive moving average models of unknown order. Biometrika 71, 599–607.

White, H., 1980. A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. Econometrica 48, 817–838.

Yamamoto, R., 2011. Order aggressiveness, pre-trade transparency, and long memory in an order-driven market. Journal of Economic Dynamics & Control 35 (11), 1938–1963.